

Empirical analysis and agent-based modeling of Lithuanian parliamentary elections

Aleksejus Kononovicius

Vilnius University, Institute of Theoretical Physics and Astronomy

Abstract

In this contribution we analyze vote share distribution across the polling stations during Lithuanian parliamentary elections of 1992, 2008 and 2012. We find that the vote share distributions are rather well fitted by the Beta distributions. To reproduce this empirical observation we propose a multi-state agent-based model in which all agents choose political party to support. The agents are allowed to change their decisions on their own as well as due to the linear recruitment mechanism. We use the simple model to reproduce vote share distributions of the 1992 election. We discuss extensions needed to reproduce vote share distributions of the other elections.

1 Introduction

While any individual's vote is equally important to determine the outcome of election, the probability that any single vote would be decisive is extremely small. In this context a rational choice would not to vote as utility is small and there is at least a minor associated cost. Yet this simplistic context may be extended to provide sound reasons to vote: to show your support for the democracy [1,2], to avoid the risk of regret [3], there might be also a social cost for not voting [4]. Among these works a few early game theoretic approaches [1,5] seemed rather promising, but further research have shown that game theoretic model of election with pure Nash equilibrium, and even mixed Nash equilibrium, might be impossible unless certain specific conditions are met [6–8]. But people are rarely ideally rational as well as they rarely face problems which might be solved by purely rational thinking [9,10].

Such context allows to consider the modeling of the voting behavior from the psychological perspective [11–18]. The main drawback of the psychologically motivated models is the increased complexity, which make models harder to implement and interpret, as well as increased number of parameters, which could allow to fit almost anything. Although recently a psychologically motivated model was quite successfully used to discuss political climate in Poland as well as to forecast Polish election of 2015 [15,17].

A different perspective comes from physics, which is neatly summarized by the Boltzmann's molecular chaos hypothesis [19]:

The molecules are like so many individuals, having the most various states of motion, and the properties of gases only remain unaltered because the number of these molecules which on the average have a given state of motion is constant.

During the last three decades physicists have approached social and economic systems from this perspective, looking for universal laws and important statistical patterns yet keeping the model as simple as it is possible. This effort by a numerous more or less prominent physicists became what is now known as econophysics and sociophysics [20–25]. Opinion dynamics, and voting behavior as a proxy of opinion, is still one of the major topics in sociophysics [23–26].

This paper contributes to understanding and describing voting dynamics from couple of different point of views. First of all we consider data sets from the Lithuanian parliamentary elections (Lithuania being young democracy), while most of the existing sociophysics literature have considered different elections held in a mature democracies, such as Brazil, England, Germany, France, Finland, Norway, Switzerland and numerous others [27–31]. In political and sociological literature one could also find numerous previous approaches to Lithuanian parliamentary elections, e.g., [32–36]. Yet most of these approaches had quite different, broader, perspective discussing general electoral trends facing varying social, demographic and economic changes. For such discussion highly aggregated data is sufficient, while for our purpose we need to consider the data at smallest scale available (polling station).

Another key contribution of this paper is a simple agent-based model, which we use to explain patterns uncovered in our analysis. The proposed model is built upon a two-state herding model originally proposed by Kirman in [37]. In the recent years the two-state herding model was quite frequently and rather successfully applied to model various statistical patterns observed in empirical data of the financial markets [38–44]. Here we extend the two-state herding model to allow multiple states and switching between them. We discuss the similarity between the proposed multi-state model and the well known Voter model [45–50]. Our approach is unique in a sense that we consider reproducing vote share distribution of elections in Lithuania. In previous literature there was but a single attempt to model, and predict, total vote share received by the parties in Lithuanian parliamentary elections using regressive model, see [51]. A major part of the previous sociophysical papers have ignored vote share distributions, likely due to belief that party vote share reflects electoral sensitivity to the policies promoted by the parties and less to the endogenous interactions between the voters (similar argument is given in [27]). To some extent this belief is supported by game theoretic models, see [1]. Yet there were some notable sociophysical modeling approaches considering agents making a binary choices, e.g., voting for or against certain proposals in a referendum [26]. It is worth to note that recently the binary models considered in [26] were used to construct a simple financial market model [52] (while our approach in this paper is the opposite). The other statistical patterns arising during various type of elections were considered and modeled, e.g., a branching process model was proposed to reproduce individual, nominated via open party list, vote share distribution [27], a network model was used to explain how individuals decide to participate in municipal elections [30], a diffusive model for the election turn-out was proposed in [31]. One of the more similar approaches was taken by [53], in which a generative model was proposed to reproduce the party vote share rank-size distributions. Another similar approach [50] considered the party vote share probability distribution and reproduced it using the multi-state mean-field Voter model.

This paper is organized as follows. In Section 2 we discuss Lithuanian parliamentary election system as well as carry out empirical analysis. Next we will briefly introduce Kirman’s herding model and extend the model to account for the multiple states (see Section 3). Afterwards, in Section 4, we will apply the multi-state agent-based model to reproduce statistical patterns uncovered in our analysis. Finally we will gather final remarks in Section 5.

2 Empirical analysis of Lithuanian parliamentary elections

Let us start the analysis by discussing the mixed parliamentary election system used in Lithuania. Lithuanian parliamentary elections are being held each 4 years. During the election all of the 141 parliamentary seats are distributed using two-tier voting system. Namely, each voter casts voter for an open party list (a person may vote for up to 5 individuals on the that list) and also for a representative of his electoral district. As a result 71 seats are allocated to the district representatives, while the other 70 seats are distributed proportionally among the parties which received more than 5% of votes for their list. Individuals on the list are ranked by popularity according to the optionally cast votes. Each electoral district (there are 71 of them) has multiple polling stations (the number varies over different elections), which correspond to further territorial subdivision of the electoral districts. Each registered voter is assigned to a single polling station based on where he lives. Different polling

stations might have very different number of registered voters assigned to them – smallest polling stations might have 100 registered voters, while largest up to 7000.

In this paper we consider only voting for the open party lists, for the parties themselves, in each of the local polling stations. Namely we ignore votes cast in polling stations abroad or votes cast by post. We also ignore ranking statistics of the individuals on the party list (considered in [27]), voting for the representative of electoral district (considered in [50,53]) and participation rates (modeling and analysis of which was considered in [29,31]). In our analysis that follows we consider only parties which have total vote share larger than 5% (at least one seat obtained according to popular vote), combining all of the other parties into a single “Other” party.

In this paper we analyze three data sets from the Lithuanian parliamentary elections of 1992, 2008 and 2012. All original data sets are publicly available from Central Electoral Commission of the Republic of Lithuania (on <https://www.rinkejopuslapis.lt/ataskaitu-formavimas> website). We have downloaded the original data sets from the website on August 31, 2016. During the initial analysis we have found some small inconsistencies within the original data. The original data set of the 1992 election had seven polling stations with incorrect total vote counts. We have identified three pairs of polling stations which were, most likely, swapped among themselves – the number of missing votes in one polling station matched the number of surplus votes in the other. While we have dealt with the odd polling station by simply adjusting the total vote count to match a sum of votes given to each of the parties in that polling station. We have also found that data from 51 (out of 2034) polling stations is missing (the data was filled with zeros) in the original 2008 election data set. We have not identified any issues with the data from the 2012 election. These minor inconsistencies would not impact the overall result of the elections nor the results reported in this section. The cleaned up version of the data sets are available online on <https://github.com/akononovicius/lithuanian-parliamentary-election-data>.

In the analysis that follows we consider party vote share in each polling station. The vote share, v_{ij} , is defined as total number votes cast for the party V_{ij} divided by the total number of votes cast in that polling station:

$$v_{ij} = \frac{V_{ij}}{\sum_{k=1}^P V_{kj}}, \quad (1)$$

here index i varies over distinct parties (P is a total number of parties participating in election) and index j varies over distinct polling stations. We consider probability and rank-size distribution of v_{ij} over all polling stations during the same election. The probability distribution is estimated using probability density functions (abbr. PDF).

We also consider the rank-size distributions of the vote share. Originally this technique is common in applications where the data varies significantly in scale, such as word occurrence frequency [54], large earthquake magnitude [55], city sizes [56], cross country income distribution [57]. When using this technique a naturally occurring unsorted data is sorted in descending order (largest first) according to the value of relevant parameter. Afterwards the sorted data is plotted with rank on the abscissa axis and the value of the parameter on the ordinate axis. In our case we sort unsorted vote share data, v_{ij} , of each party i over polling stations j to produce \tilde{v}_{ik} , for which

$$\tilde{v}_{i1} \geq \tilde{v}_{i2} \geq \dots \geq \tilde{v}_{iM} \quad (2)$$

is true. In the above M is a total number of polling stations, so that index k represents correspond the rank. Evidently these two statistics, PDF and rank-size distribution, are inter-related, but using both of them allows to uncover certain statistical patterns.

2.1 Parliamentary election of 1992

Parliamentary election of 1992 was held in 2061 polling stations (thus we have 2061 data points). 17 parties competed in the election, but only 4 of them were able to obtain more than 5% of votes. For the sake of simplicity we will use the following abbreviations for these more popular parties: SK – “Sąjūdžio koalicija”, LSDP – “Lietuvos socialdemokratų partija”, LKDP – “Lietuvos Krikščionių demokratų partijos, Lietuvos

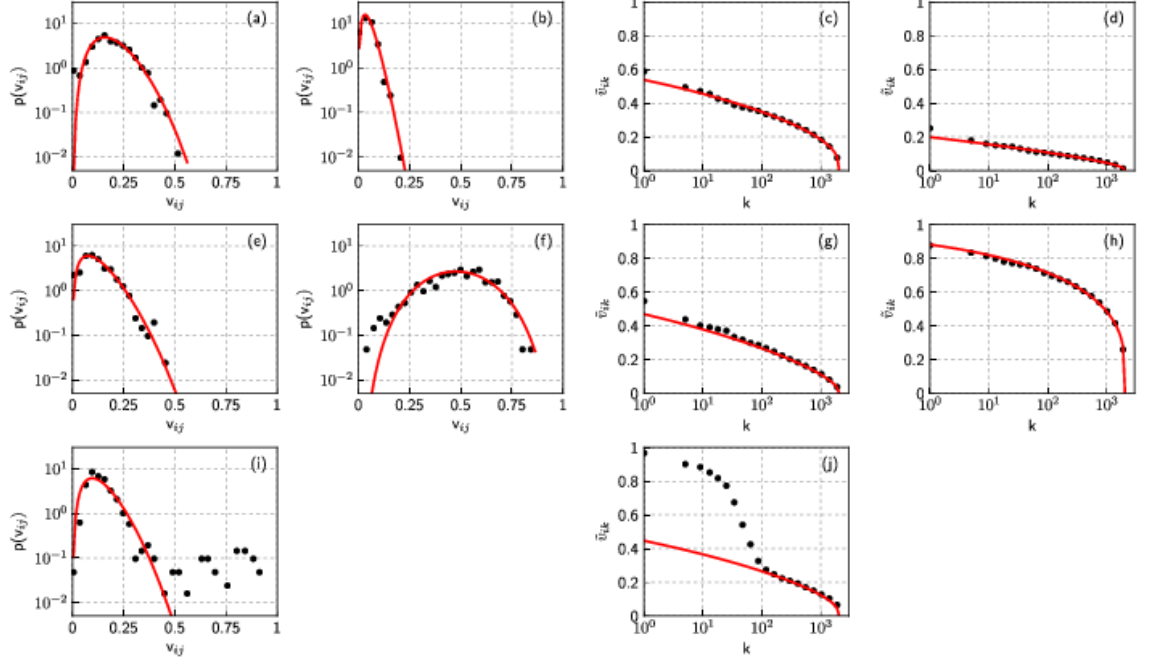


Figure 1: Vote share PDF (two left-most columns) and rank-size distribution (two right-most columns) of the most popular parties during 1992 election. The following parties were considered: SK ((a) and (c)), LSDP ((b) and (d)), LKDP ((e) and (g)), LDDP ((f) and (h)) and O ((i) and (j)). The empirical values are shown as black circles, while theoretical fits using Beta distribution are provided as red solid curves (the values of parameters are given in Table 1).

Table 1: Parameters of Beta distribution, α and β , used to fit the data in Fig. 1 as well as wellness of fit for PDFs, R_{PDF}^2 , and rank-size distributions, R_{RS}^2 .

Party	α	β	R_{PDF}^2	R_{RS}^2
SK	3.9	16.6	0.956	0.994
LSDP	2.7	51	0.935	0.953
LKDP	2.2	16	0.926	0.995
LDDP	5.7	6.1	0.907	0.998
O	3	19.4	0.895	0.854

politinių kalinių ir tremtinių sąjungos ir Lietuvos demokratų partijos jungtinis sąrašas”, LDDP – “Lietuvos demokratinė darbo partija”. Votes for all of the other parties were combined as votes for the “Other” party (abbr. O).

As you can see from Fig. 1 and Table 1 all parties with a notable exception of the “Other” party are very well fitted by assuming that data is distributed according to the Beta distribution, PDF of which is given by

$$p(v_{ij}) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)} v_{ij}^{\alpha-1} (1 - v_{ij})^{\beta-1}. \quad (3)$$

The “Other” party stands out, because it includes “Lietuvos lenkų sąjunga” party (abbr. LLS; en. Association of Poles in Lithuania). This party had heavily relied on the support of ethnic minorities, which were spatially segregated. Namely, ethnic minorities mostly live in larger cities and Vilnius County. This geographical segregation could easily translate into the segregation observed in the voting data.

In Fig. 2 we confirm this intuition by splitting LLS away from the “Other” party. After the split rank-size distribution of the “Other” party is well approximated by Beta distribution with parameters $\alpha = 4.9$, $\beta = 35.4$ ($R_{RS}^2 = 0.992$). To provide a good fit for the LLS rank-size distribution one may assume that underlying data actually follows a mixture of two Beta distributions: one (95% of points) with parameters $\alpha_1 = 0.08$, $\beta_1 = 10$ and the other (5% of points) with parameters $\alpha_2 = 1.22$, $\beta_2 = 1.37$.

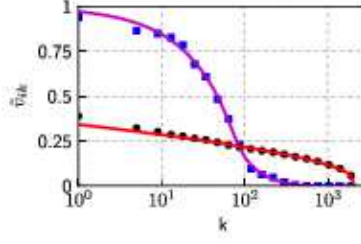


Figure 2: Rank-size distributions of LLS (blue squares) and the other parties in the “Other” party (black circles). The empirical values are shown as black circles, while solid curves provide fits by assuming that: $v_{ij} \sim \mathcal{Be}(4.9, 35.4)$ ($R_{RS}^2 = 0.992$; red curve), $v_{ij} \sim 0.95 \cdot \mathcal{Be}(0.08, 10) + 0.05 \cdot \mathcal{Be}(1.22, 1.37)$ ($R_{RS}^2 = 0.997$; magenta curve).

Table 2: Parameters of Beta distribution used to fit the data in Fig. 3 as well as wellness of fit for PDFs, R_{PDF}^2 , and rank-size distributions, R_{RS}^2 .

Party	α_1	β_1	c	α_2	β_2	R_{PDF}^2	R_{RS}^2
LSDP	3.9	31.7	0.15	4.3	12.9	0.968	0.999
TS-LKD	3.7	16.8	0	–	–	0.915	0.999
TPP	5.1	27.8	0	–	–	0.884	0.992
DP	3	30.3	0.09	3.2	7.7	0.942	0.992
LRLS	2.7	67.9	0.54	0.6	12.6	0.986	0.994
TT	7.6	59.5	0.42	1.8	9.4	0.987	0.992
LiCS	0.98	23.5	0	–	–	0.955	0.993
O	6.6	30.4	0.15	1.2	1.6	0.944	0.995

Party vote share rank-size distributions were previously considered by Fenner and others in [53]. Unlike here Fenner and others assumed that data is distributed according Weibull distribution, they have obtained rather good fits for the UK election data. Yet fits obtained here, assuming Beta distribution, are also rather good. We believe that Beta distribution is superior for this purpose from the theoretical point of view. Namely, Beta distribution has reasonable support, probabilities are defined for $v \in [0; 1]$, while Weibull distribution needs to be arbitrary truncated, as probabilities are defined for $v \in [0; +\infty)$. Interestingly Fenner and others also use a mixture distribution (of two Weibull distributions) to fit the UK election data. Similar observations were also made when studying Brazilian presidential election data [58].

2.2 Parliamentary election of 2008

Parliamentary election of 2008 was held in 2034 polling stations, but with the data from 51 polling station missing we have only 1983 points in our data set. 16 parties had participated in the election, with 7 of them being able to obtain more than 5% of votes. For the sake of simplicity we will use the following abbreviations for them: LSDP – “Lietuvos socialdemokratų partija” (formed by LSDP and LDDP, which participated in 1992 election), TS-LKD – “Tėvynės sąjunga – Lietuvos krikščionys demokratai” (could be considered to be a successor of the SK and LKDP, which participated in 1992 election), TPP – “Tautos prisikėlimo partija”, DP – “Koalicija Darbo partija + jaunimas”, LRLS – “Lietuvos Respublikos liberalų sąjūdis”, TT – “Partija Tvarka ir teisingumas”, LiCS – “Liberalų ir centro sąjunga”. Votes for all of the other parties were combined as votes for the “Other” party (abbr. O).

As is evident from Fig. 3 in 2008 parliamentary election voting statistics of the most of the parties are well described by the mixture of two Beta distributions. Although now there is no clear cut explanation for this observation, we could conjecture that this represents social and/or economic segregation of voters (e.g., party favoring higher income voters, who live in the richer districts of the larger cities, while losing votes in rural areas or areas populated by voters with smaller income).

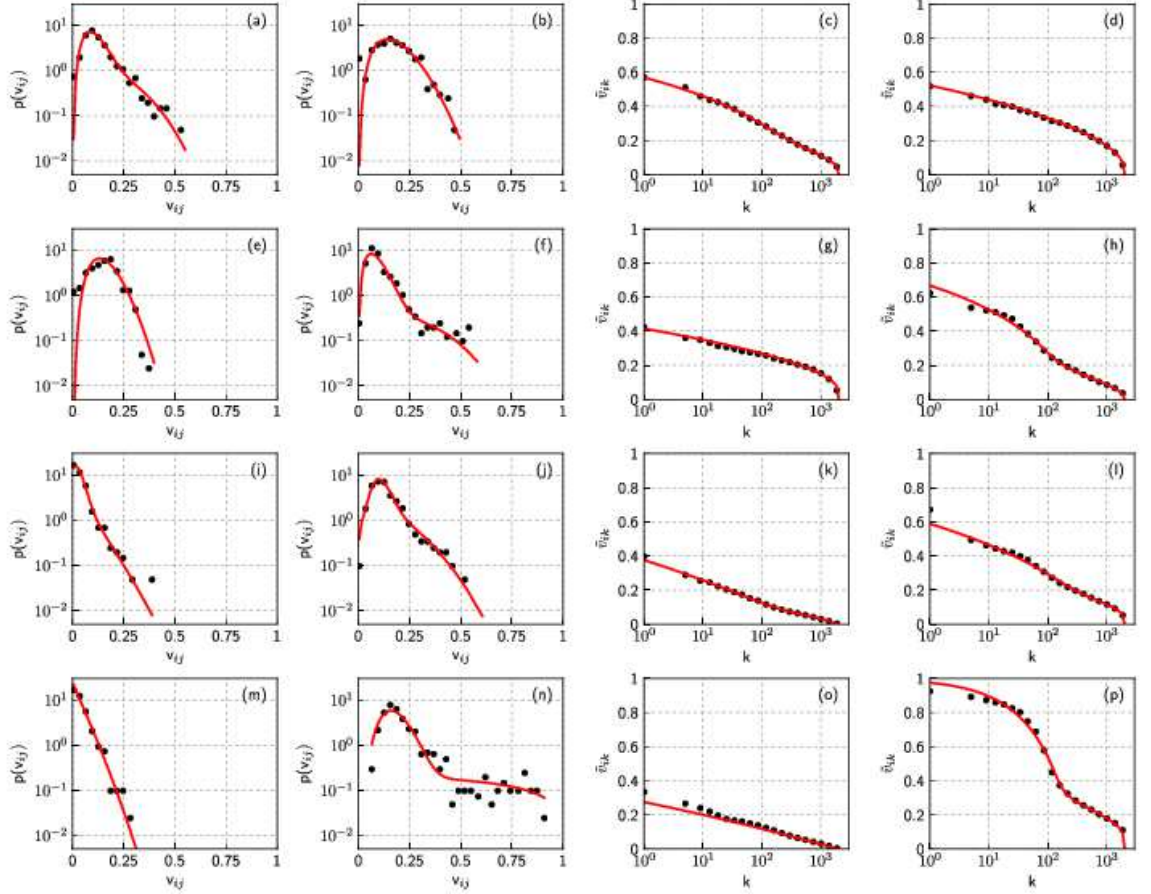


Figure 3: Vote share PDF (two left-most columns) and rank-size distribution (two right-most columns) of the most popular parties during 2008 election. The following parties were considered: LSDP ((a) and (c)), TS-LKD ((b) and (d)), TPP ((e) and (g)), DP ((f) and (h)), LRLS ((i) and (k)), TT ((j) and (l)), LiCS ((m) and (o)) and O ((n) and (p)). The empirical values are shown as black circles, while theoretical fits using Beta distribution are provided as red solid curves (the values of parameters are given in Table 2).

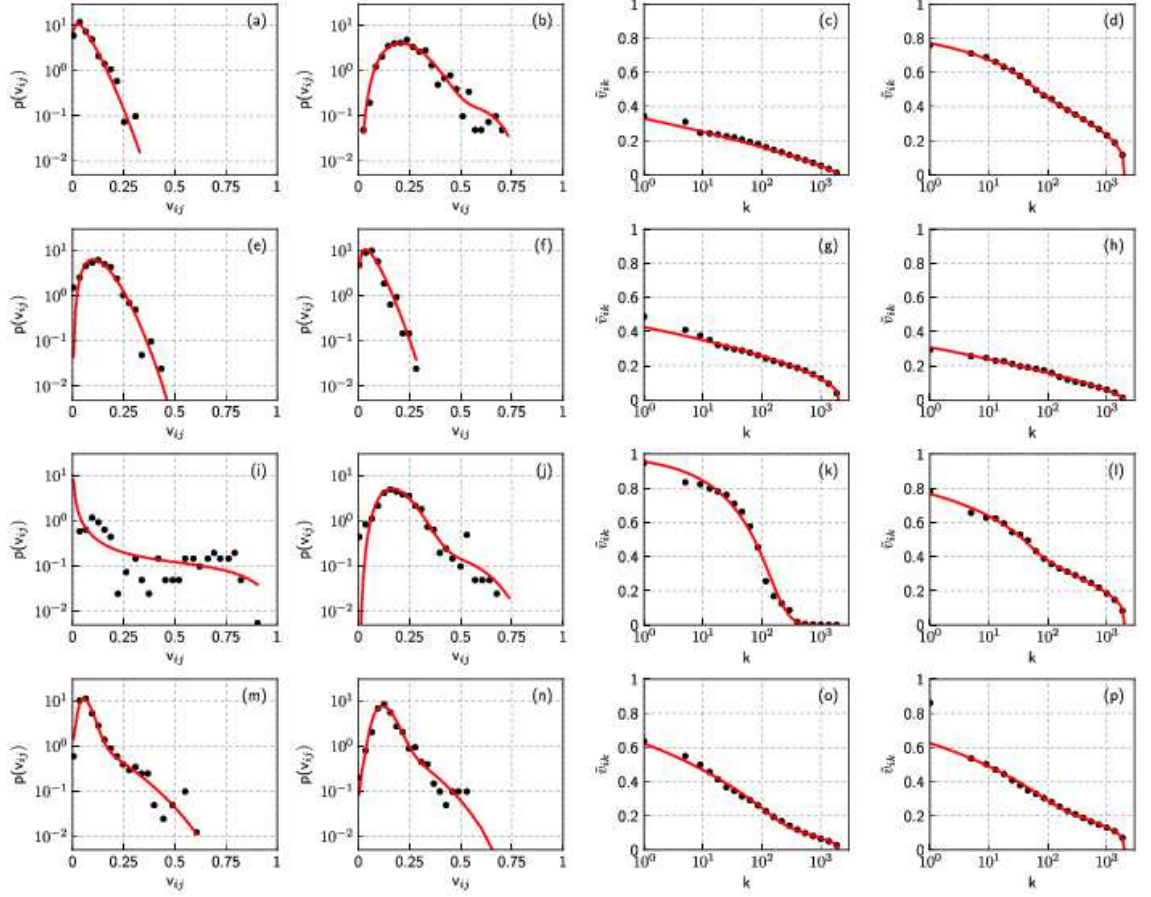


Figure 4: Vote share PDF (two left-most columns) and rank-size distribution (two right-most columns) of the most popular parties during 2012 election. The following parties were considered: LRLS ((a) and (c)), DP ((b) and (d)), TS-LKD ((e) and (g)), DK ((f) and (h)), LLRA ((i) and (k)), LSDP ((j) and (l)), TT ((m) and (o)) and O ((n) and (p)). The empirical values are shown as black circles, while theoretical fits using Beta distribution are provided as red solid curves (the values of parameters are given in Table 3).

2.3 Parliamentary election of 2012

Parliamentary election of 2012 was held in 2017 polling stations (thus we have 2017 data points). 18 parties had participated in the election, with 7 of them being able to obtain more than 5% of votes. For the sake of simplicity we will use the following abbreviations for them: LRLS – “Lietuvos Respublikos liberalų sąjūdis”, DP – “Darbo partija”, TS-LKD – “Tėvynės sąjunga – Lietuvos krikščionys demokratai”, DK – “Drąsos kelias politinė partija”, LLRA – “Lietuvos lenkų rinkimų akcija”, LSDP – “Lietuvos socialdemokratų partija”, TT – “Partija Tvarka ir teisingumas”. Matching abbreviations indicate the same (or mostly the same) parties as in 2008 election. Votes for all of the other parties were combined as votes for the “Other” party (abbr. O).

Once again, as well as in 2008 parliamentary election data set, it is evident that the voting statistics of the most of the parties in 2012 parliamentary are also well described by the mixture of two Beta distributions (see Fig. 4).

3 Multi-state agent-based model for the party vote share

In this Section we propose a simple multi-state agent-based model purpose of which is to describe opinion dynamics within a small non-specific geographic area assigned to a single polling station. Unlike in some [11–18] previous approaches we completely ignore actual psychologically motivated mechanisms and only focus on reproducing the empirical statistics discussed in the previous section.

Table 3: Parameters of Beta distribution used to fit the data in Fig. 4 as well as wellness of fit for PDFs, R_{PDF}^2 , and rank-size distributions, R_{RS}^2 .

Party	α_1	β_1	c	α_2	β_2	R_{PDF}^2	R_{RS}^2
LRLS	1.5	22	0	–	–	0.969	0.994
DP	4.5	14.5	0.03	15.3	11.1	0.973	0.999
TS-LKD	3.4	22	0	–	–	0.938	0.982
DK	1.9	26	0	–	–	0.946	0.994
LLRA	0.06	1.9	0.05	2.1	1.9	0.857	0.990
LSDP	5	21.4	0.05	5.5	6.6	0.948	0.997
TT	4.6	62	0.29	1.1	6.7	0.963	0.995
O	7	45.8	0.23	2	8	0.966	0.970

Originally in [37] it was noted that biologists and economists observe similar behavioral patterns. Namely, in entomological experiments [59, 60] it was observed that ants prefer using the same foraging paths as most of the other ants. While it is known that people also show interest in things which are more popular among their peers, e.g., restaurants, books, movies and others [61–63], regardless of their actual quality. In [37] a simple two-state model was proposed to explain such observations. This model assumed that agents switch their state idiosyncratically, act according to the perceived attractiveness of the available options, as well as are influenced by the linear recruitment mechanism. Mathematically this model was formulated using one step transition probabilities. In the contemporary iterations of the original model the following mathematical form of the transition probabilities is used [38–41]:

$$P(X \rightarrow X + 1) = (N - X)(\sigma_1 + hX) \Delta t, \quad (4)$$

$$P(X \rightarrow X - 1) = X[\sigma_2 + h(N - X)] \Delta t, \quad (5)$$

here N is a total number of agents acting in the system, X – total number of agents occupying the first state (consequently there are $N - X$ agents occupying the second state), σ_i – perceived state attractiveness parameters, h – recruitment efficiency parameter and Δt – relatively short time step. The time step, Δt , is chosen to be as small so that during the time step only a single agent would be likely change his state, otherwise agent-based dynamics would not be well approximated by the one step transition probabilities. In the scope of this paper h parameter is not relevant as its main use is alignment of model time scale to the real time scale [38–41]:

$$P(X \rightarrow X + 1) = (N - X)(\varepsilon_1 + X) \Delta t_s, \quad (6)$$

$$P(X \rightarrow X - 1) = X[\varepsilon_2 + (N - X)] \Delta t_s, \quad (7)$$

here $\varepsilon_i = \frac{\sigma_i}{h}$ is relative attractiveness and $t_s = ht$ is a rescaled time. This two-state model may be given aggregate macroscopic description, using Fokker–Planck or stochastic differential equations, using birth–death process formalism [39, 41, 64, 65]. From the stochastic differential equation describing the dynamics of this model it is rather straightforward to conclude that the distribution of $x = \frac{X}{N}$ is $\mathcal{Be}(\varepsilon_1, \varepsilon_2)$ [38, 39].

As in typical parliamentary election there are more than two competitors, we need to reconsider one step transition probabilities. From the conservation of total number of agents N , we have:

$$P(X_i \rightarrow X_i \pm 1) = \sum_{j \neq i} P(X_i \rightarrow X_i \pm 1, X_j \rightarrow X_j \mp 1). \quad (8)$$

In general case expanding the right hand side in the same form as Eqs. (4) and (5):

$$P(X_i \rightarrow X_i + 1) = \sum_{j \neq i} X_j (\sigma_{ji} + h_{ji} X_i) \Delta t, \quad (9)$$

$$P(X_i \rightarrow X_i - 1) = X_i \sum_{j \neq i} [\sigma_{ij} + h_{ij} X_j] \Delta t, \quad (10)$$

These one step transition probabilities for X_i will depend not only on X_i (as is with the two-state model), but also on all other X_j (where $j \neq i$). One needs to assume that $\sigma_{ij} = \sigma_j$ and $h_{ij} = h$ (where $j \neq i$) in order to avoid cumbersome dependence on all other X_j . The first assumption, $\sigma_{ij} = \sigma_j$, means that the perceived attractiveness of any party is independent of by whom it is perceived. While the second assumption, $h_{ij} = h$, means that recruitment is symmetric and independent of supporters of which parties interact. Note that the both of these assumptions contrast with the assumptions underlying the bounded confidence model [13, 14]. Yet these assumptions are needed so that we could be sure that $x_i = \frac{X_i}{N}$ is distributed according to Beta distribution. Using these assumptions we can simplify the one step transition probabilities:

$$P(X_i \rightarrow X_i + 1) = (N - X_i)(\varepsilon_i + X_i)\Delta t_s, \quad (11)$$

$$P(X_i \rightarrow X_i - 1) = X_i(\varepsilon_{-i} + N - X_i)\Delta t_s, \quad (12)$$

here ε_{-i} is the total attractiveness of the competitors of party i , $\varepsilon_{-i} = \sum_{j \neq i} \varepsilon_j$. By the analogy with the two state model it should be evident that

$$x_i \sim \mathcal{Be}(\varepsilon_i, \varepsilon_{-i}). \quad (13)$$

Unlike the two-state model, it seems impossible to provide a general aggregated macroscopic description of the N -state model. In [66] three-state model was considered and given aggregated macroscopic description, by a system of two stochastic differential equations, yet it was done under specific conditions.

The one step transition probabilities, while consider agent behavior, are still aggregate description of agent level dynamics. So some discussion on what do the Eqs. (11) and (12) represent is relevant. Selecting one random agent, per time step, and setting his switching probability to $\varepsilon_{-i}\Delta t_s$ gives us idiosyncratic behavior term, $X_i\varepsilon_{-i}\Delta t_s$. While selecting another random agent and, if both agents vote for different parties, allowing the first agent to copy the second agent's voting behavior gives us the recruitment term, $X_i(N - X_i)\Delta t_s$. This description could be further generalized to allow the model to be run on the randomly generated networks [67, 68]. This agent-based algorithm might be seen to be a special case of the well known Voter model [23, 45–47, 49, 50].

4 Modeling of the parliamentary elections

In this Section we apply the model presented in previous Section to reproduce PDFs and rank-size distributions observed during the parliamentary elections of 1992. We consider only the simplest case by ignoring the “Other” party (removing all small parties from our data set) and thus removing the segregation. We do not consider segregated data, full data of the 1992 parliamentary election or data of the 2008 and 2012 parliamentary elections, as to account for this one needs to take a rather complicated approach. It would mostly impossible to guess the correct partition of the polling stations to separate data sets, which could be modeled separately using a single parameter set. Although in general such approach would be conceivable if a detailed spatial opinion polling data or detailed socio-demographic data would be available. This would effectively lead to the spatial voting model of Lithuania.

In Fig. 5 we compare the vote share statistics generated by the proposed model, Eqs. (11) and (12), and empirical vote share statistics of the parliamentary election of 1992. In this comparison we have ignored the “Other” party as its vote share statistics exhibits notable segregation, while the separately considered the more popular parties appear not to be notably affected by the segregation (see the discussion in Section 2.1). Yet we cannot use the previously empirically estimated Beta distribution parameters, by assuming $\alpha_i = \varepsilon_i$ and $\beta_i = \varepsilon_{-i}$, as model parameters, because an important model implication, $\sum_{l \neq i} \alpha_l = \beta_i$, doesn't hold for the empirical data. Yet one may obtain the parameter values by fitting the empirical data with the model implication in mind.

As you can see from Fig. 5 as well as Table 4 the proposed model excellently fits three of the four parties. While for the LSDP the fit is not as good as one might expect. Note that model overestimates success of LSDP (in higher ranked polling stations it should have had larger vote share) and underestimates success of LDDP (in

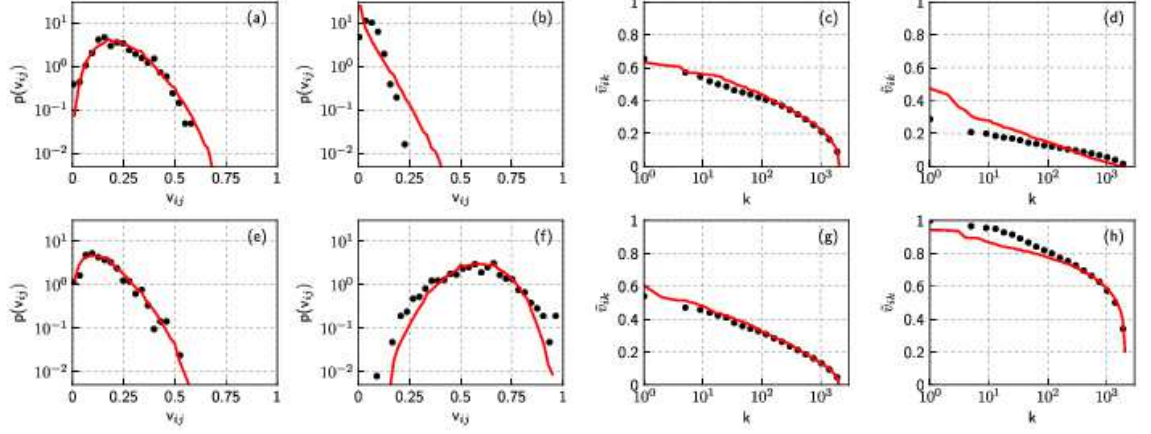


Figure 5: Vote share PDF (two left-most columns) and rank-size distribution (two right-most columns) of the most popular parties during 1992 election. Only the following parties were considered: SK ((a) and (c)), LSDP ((b) and (d)), LKDP ((e) and (g)) and LDDP ((f) and (h)). The empirical values are shown as black circles, while red solid curves represent data numerically generated by the proposed model, Eqs. (11) and (12). Model parameters are given in Table 4.

Table 4: Parameters of the proposed model used to reproduce the vote share statistics of parliamentary election of 1992 as well as wellness of fits for respective statistical properties.

Party	ε_i	R_{PDF}^2	R_{RS}^2
SK	3.52	0.951	0.997
LSDP	1.13	0.705	0.881
LKDP	2.27	0.952	0.998
LDDP	8.87	0.904	0.973

higher ranked polling stations it should have had smaller vote share). Thus it is likely that LSDP had small perceived chance to win the election, thus the voters who would consider voting for LSDP actually voted for another left-wing party, which had better perceived chance at winning (LDDP).

We can check this intuition by violating our assumption that perceived attractiveness should not depend on the current state of the agent, namely instead of ε_i we now have ε_{ji} . Previously this assumption was needed to ensure that vote share is distributed according to Beta distribution. Let us keep the previous relation $\varepsilon_{ji} = \varepsilon_i$ intact with a single exception when j corresponds to LSDP and i corresponds to LDDP (the numeric indices are assigned according to Table 4). This gives us the following matrix of ε_{ji} values:

$$\varepsilon = \begin{pmatrix} 0 & 1.13 & 2.27 & 8.87 \\ 3.52 & 0 & 2.27 & 22 \\ 3.52 & 1.13 & 0 & 8.87 \\ 3.52 & 1.13 & 2.27 & 0 \end{pmatrix}. \quad (14)$$

Note that the diagonal elements are set to zero, as it is not possible to switch to the state the agent is already in. As you can see in Fig. 6, the model fit provided for LSDP has improved significantly by making this small change.

5 Conclusions

In this paper we have considered vote share PDFs and rank-size distributions of the parliamentary elections held in Lithuania. Namely, we have considered three data sets of the 1992, 2008 and 2012 parliamentary elections. The empirical statistical properties were rather well fitted by assuming that the data is distributed according to Beta distribution or a mixture of two Beta distributions. In our literature review we have found

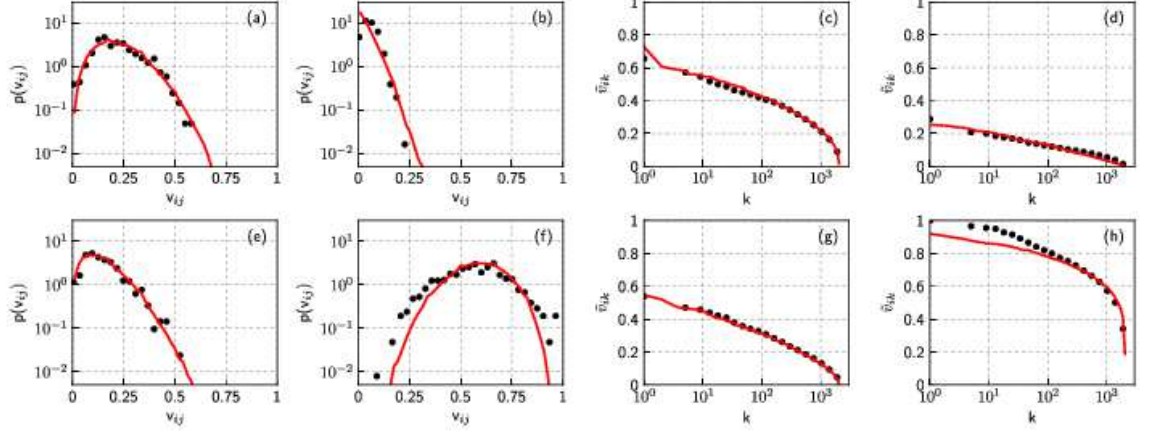


Figure 6: Vote share PDF (two left-most columns) and rank-size distribution (two right-most columns) of the most popular parties during 1992 election. Only the following parties were considered: SK ((a) and (c)), LSDP ((b) and (d)), LKDP ((e) and (g)) and LDDP ((f) and (h)). The empirical values are shown as black circles, while red solid curves represent data numerically generated by the proposed model with attractiveness dependent on both current agent state and the perceived state. Model parameters are given by Eq. (14).

that [50, 53, 58] have reported somewhat similar results. In [53, 58] it was reported that the empirical data is rather well fitted by Weibull distribution or a mixture of Weibull distributions. We argue that Beta distribution is more suitable as it has correct support (probabilities are defined for $v \in [0; 1]$; although Weibull distribution, originally $v \in [0; +\infty)$, could be arbitrary truncated) and it arises from a simple easily tractable agent-based model. From our empirical analysis it follows that the mixture of distributions is needed to fit the data if there is underlying spatial segregation of voters. [50] does report that the empirical data is rather well fitted by Beta distribution and have used the Voter model to provide theoretical background for this observation. Yet [50] had not reported observation of the segregation patterns.

Having in mind the stark difference between the psychologically motivated models, such as bounded confidence model [13, 14], we would like to point out that the observed statistical patterns as well as applicability of the model could arise due to numerous reasons. One of the alternative possibilities would be the people mobility patterns. In the proposed model a single agent switching from supporting one party to supporting another party, could also represent one agent moving away from the modeled geographic location, due to social or economic reasons, and another agent, holding different political views, moving in. Similar suggestion was also raised in [49].

In the nearest future we will consider spatial modeling of Lithuanian parliamentary elections. Another possible approach, with forecasting possibility, could be considering temporal regressive model for the attractiveness parameters of the proposed model, ε_i , as well as estimation of agent interaction rates, h .

Acknowledgements

The author would like to thank Ainė Ramonaitė for suggesting the idea to analyze Lithuanian parliamentary election data as well as many useful remarks. The author would also like to acknowledge valuable discussions and general interest in this venture show by colleagues at VU ITPA Julius Ruseckas and Vygintas Gontis.

References

- [1] A. Downs, “An economic theory of political action in a democracy,” *Journal of Political Economy*, vol. 65, no. 2, pp. 135–150, 1957.

- [2] W. H. Riker and P. C. Ordeshook, “A theory of the calculus of voting,” *The American Political Science Review*, vol. 62, no. 1, pp. 25–42, 1968.
- [3] J. A. Ferejohn and M. P. Fiorina, “The paradox of not voting: A decision theoretic analysis,” *American Political Science Review*, vol. 68, no. 2, pp. 525–536, 1974.
- [4] C. J. Uhlaner, “Rational turnout: The neglected role of groups,” *American Journal of Political Science*, vol. 33, no. 2, pp. 390–422, 1989.
- [5] H. Hotelling, “Stability in competition,” *The Economic Journal*, vol. 39, no. 153, pp. 41–57, 1929.
- [6] C. R. Plott, “A notion of equilibrium and its possibility under majority rule,” *The American Economic Review*, vol. 57, no. 4, pp. 787–806, 1967.
- [7] R. D. McKelvey, “Intransitivities in multidimensional voting models and some implications for agenda control,” *Journal of Economic Theory*, vol. 12, no. 3, pp. 472–482, 1976.
- [8] S. Ansolabehere and J. M. Snyder, “Valence politics and equilibrium in spatial election models,” *Public Choice*, vol. 103, no. 3/4, pp. 327–336, 2000.
- [9] W. B. Arthur, “Inductive reasoning and bounded rationality,” *American Economic Review*, vol. 84, pp. 406–411, 1994.
- [10] G. Akerlof and J. Shiller, *Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism*. Princeton University Press, 2009.
- [11] A. Nowak, J. Szamrej, and B. Latane, “From private attitude to public opinion: A dynamic theory of social impact,” *Psychological Review*, vol. 97, no. 3, pp. 362–376, 1990.
- [12] R. Axelrod, “The dissemination of culture a model with local convergence and global polarization,” *Journal of Conflict Resolution*, vol. 41, no. 2, pp. 203–226, 1997.
- [13] R. Hegselmann and U. Krause, “Opinion dynamics and bounded confidence: models, analysis and simulation,” *Journal of Artificial Societies and Social Simulation*, vol. 5, no. 3, p. 2, 2002.
- [14] G. Deffuant, “Comparing extremism propagation patterns in continuous opinion models,” *Journal of Artificial Societies and Social Simulation*, vol. 9, no. 3, p. 8, 2006.
- [15] P. Sobkowicz, “Discrete model of opinion changes using knowledge and emotions as control variables,” *PLoS ONE*, vol. 7, no. 9, p. e44489, 2012.
- [16] D. G. Lilleker, *Modelling Political Cognition*. London: Palgrave Macmillan UK, 2014, pp. 198–205.
- [17] P. Sobkowicz, “Quantitative agent based model of opinion dynamics: Polish elections of 2015,” *PLoS ONE*, vol. 11, no. 5, p. e0155098, 2016.
- [18] P. Duggins, “A psychologically-motivated model of opinion change with applications to american politics,” *Journal of Artificial Societies and Social Simulation*, vol. 20, no. 1, p. 13, 2017.
- [19] P. Ball, “The physical modelling of society: A historical perspective,” *Physica A*, vol. 314, pp. 1–14, 2002.
- [20] S. Galam, “Majority rule, hierarchical structures, and democratic totalitarianism: A statistical approach,” *Journal of Mathematical Psychology*, vol. 30, no. 4, pp. 426 – 434, 1986.
- [21] R. N. Mantegna and H. E. Stanley, *Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge University Press, 2000.
- [22] B. K. Chakrabarti, A. Chakraborti, and A. Chatterjee, *Econophysics and Sociophysics - Trends and Perspectives*. Wiley, 2006.

- [23] C. Castellano, S. Fortunato, and V. Loreto, “Statistical physics of social dynamics,” *Reviews of Modern Physics*, vol. 81, pp. 591–646, 2009.
- [24] D. Stauffer, “A biased review of sociophysics,” *Journal of Statistical Physics*, 2013.
- [25] F. Abergel, H. Aoyama, B. Chakrabarti, A. Chakraborti, N. Deo, D. Raina, and I. Vodenska, Eds., *Econophysics and Sociophysics: Recent Progress and Future Directions*. Springer, 2017.
- [26] S. Galam, “Sociophysics: A review of galam models,” *International Journal of Modern Physics C*, vol. 19, no. 03, p. 409, 2008.
- [27] S. Fortunato and C. Castellano, “Scaling and universality in proportional elections,” *Physical Review Letters*, vol. 99, p. 138701, 2007.
- [28] C. A. Andresen, H. F. Hansen, A. Hansen, G. L. Vasconcelos, and J. S. Andrade Jr., “Correlations between political party size and voter memory: Statistical analysis of opinion polls,” *International Journal of Modern Physics C*, vol. 19, p. 1647, 2008.
- [29] C. Borghesi and J.-P. Bouchaud, “Spatial correlations in vote statistics: a diffusive field model for decision-making,” *European Physical Journal B*, vol. 75, no. 3, pp. 395–404, 2010.
- [30] M. C. Mantovani, H. V. Ribeiro, M. V. Moro, S. Picoli Jr., and R. S. Mendes, “Scaling laws and universality in the choice of election candidates,” *EPL*, vol. 96, no. 4, p. 48001, 2011.
- [31] C. Borghesi, J.-C. Raynal, and J. P. Bouchaud, “Election turnout statistics in many countries: similarities, differences, and a diffusive field model for decision-making,” *PLoS ONE*, vol. 7, p. e36289, 2012.
- [32] A. Krupavicius, “The lithuanian parliamentary elections of 1996,” *Electoral Studies*, vol. 16, no. 4, pp. 541 – 549, 1997.
- [33] M. Degutis, “How lithuanian voters decide: Reasons behind the party choice,” *Lithuanian Political Science Yearbook*, pp. 69–111, 2000.
- [34] M. Kreuzer and V. Pettai, “Patterns of political instability: Affiliation patterns of politicians and voters in post-communist estonia, latvia, and lithuania,” *Studies in Comparative International Development*, vol. 38, pp. 76–98, 2003.
- [35] M. Jurkynas, “Emerging cleavages in new democracies: The case of lithuania,” *Journal of Baltic Studies*, vol. 35, no. 3, pp. 278–296, 2004.
- [36] A. Ramonaite, *The Development of the Lithuanian Party System: From Stability to Perturbation*. Routledge, 2006, pp. 69–88.
- [37] A. P. Kirman, “Ants, rationality and recruitment,” *Quarterly Journal of Economics*, vol. 108, pp. 137–156, 1993.
- [38] S. Alfarano, T. Lux, and F. Wagner, “Estimation of agent-based models: The case of an asymmetric herding model,” *Computational Economics*, vol. 26, no. 1, pp. 19–49, 2005.
- [39] —, “Time variation of higher moments in a financial market with heterogeneous agents: An analytical approach,” *Journal of Economic Dynamics and Control*, vol. 32, pp. 101–136, 2008.
- [40] V. Alfi, M. Cristelli, L. Pietronero, and A. Zaccaria, “Minimal agent based model for financial markets i: Origin and self-organization of stylized facts,” *European Physics Journal B*, vol. 67, no. 3, pp. 385–397, 2009.
- [41] A. Kononovicius and V. Gontis, “Agent based reasoning for the non-linear stochastic models of long-range memory,” *Physica A*, vol. 391, no. 4, pp. 1309–1314, 2012.

- [42] S. Alfarano, M. Milakovic, and M. Raddant, “A note on institutional hierarchy and volatility in financial markets,” *The European Journal of Finance*, vol. 19, no. 6, pp. 449–465, 2013.
- [43] V. Gontis and A. Kononovicius, “Consentaneous agent-based and stochastic model of the financial markets,” *PLoS ONE*, vol. 9, no. 7, p. e102201, 2014.
- [44] V. Gontis, S. Havlin, A. Kononovicius, B. Podobnik, and H. Stanley, “Stochastic model of financial markets reproducing scaling and memory in volatility return intervals,” *Physica A*, vol. 462, pp. 1091–1102, 2016.
- [45] P. Clifford and A. Sudbury, “A model for spatial conflict,” *Biometrika*, vol. 60, pp. 581 – 588, 1973.
- [46] T. Liggett, *Stochastic Interacting Systems: Contact, Voter, and Exclusion Processes*. Springer, 1999.
- [47] S. Galam, B. Chopard, and M. Droz, “Killer geometries in competing species dynamics,” *Physica A*, vol. 314, pp. 256 – 263, 2002.
- [48] C. Castellano, D. Vilone, and A. Vespignani, “Incomplete ordering of the voter model on small-world networks,” *EPL*, vol. 63, no. 1, p. 153, 2003.
- [49] J. Fernandez-Gracia, K. Suchecki, J. J. Ramasco, M. San Miguel, and V. M. Eguiluz, “Is the voter model a model for voters?” *Physical Review Letters*, vol. 112, p. 158701, 2014.
- [50] F. Sano, M. Hisakado, and S. Mori, “Mean field voter model of election to the house of representatives in japan,” arXiv:1702.03603 [physics.soc-ph].
- [51] M. Jastramskis, “Forecasting the results of seimas elections: model of government parties’ votes,” *Parlamento studijos*, vol. 11, 2011.
- [52] S. Galam, “The invisible hand and the rational agent are behind bubbles and crashes,” *Chaos, Solitons and Fractals*, vol. 88, pp. 209 – 217, 2016.
- [53] T. Fenner, M. Levene, and G. Loizou, “A multiplicative process for generating the rank-order distribution of uk election results,” arXiv:1609.04282 [physics.soc-ph].
- [54] G. K. Zipf, *Human Behavior and the Principle of Least Effort*. Addison-Wesley, 1949.
- [55] D. Sornette, L. Knopoff, Y. Y. Kagan, and C. Vanneste, “Rank-ordering statistics of extreme events: Application to the distribution of large earthquakes,” *Journal of Geophysical Research*, vol. 101, no. B6, pp. 13 883–13 893, 1996.
- [56] X. Gabaix, “Zipf’s law for cities: An explanation,” *The Quarterly Journal of Economics*, vol. 114, no. 3, p. 739, 1999.
- [57] J. Shao, P. C. Ivanov, B. Urosevic, H. E. Stanley, and B. Podobnik, “Zipf rank approach and cross-country convergence of incomes,” *EPL*, vol. 94, p. 48001, 2011.
- [58] R. F. d. Paz, R. S. Ehlers, and J. L. Bazan, *A Weibull Mixture Model for the Votes of a Brazilian Political Party*. Springer International Publishing, 2015, pp. 229–241.
- [59] J. M. Pasteels, J. L. Deneubourg, and S. Goss, “Self-organization mechanisms in ant societies (i): Trail recruitment to newly discovered food sources,” in *From Individual to Collective Behaviour in Social Insects*, J. M. Pasteels and J. L. Deneubourg, Eds. Basel: Birkhauser, 1987, pp. 155–175.
- [60] —, “Self-organization mechanisms in ant societies (ii): Learning in foraging and division of labor,” in *From Individual to Collective Behaviour in Social Insects*, J. M. Pasteels and J. L. Deneubourg, Eds. Basel: Birkhauser, 1987, pp. 177–196.

- [61] F. M. Bass, “A new product growth model for consumer durables,” *Management Science*, vol. 15, pp. 215–227, 1969.
- [62] G. S. Becker, “A note on restaurant pricing and other examples of social influence on price,” *Journal of Political Economy*, vol. 99, pp. 1109–1116, 1991.
- [63] A. Ishii, H. Arakaki, N. Matsuda, S. Umemura, T. Urushidani, N. Yamagata, and N. Yoshida, “The ‘hit’ phenomenon: a mathematical model of human dynamics interactions as a stochastic process,” *New Journal of Physics*, vol. 14, p. 063018, 2012.
- [64] M. Aoki and H. Yoshikawa, *Reconstructing Macroeconomics: A Perspective from Statistical Physics and Combinatorial Stochastic Processes*. Cambridge University Press, 2007.
- [65] N. G. van Kampen, *Stochastic process in Physics and Chemistry*. Amsterdam: North Holland, 2007.
- [66] A. Kononovicius and V. Gontis, “Three state herding model of the financial markets,” *EPL*, vol. 101, p. 28001, 2013.
- [67] S. Alfarano and M. Milakovic, “Network structure and n-dependence in agent-based herding models,” *Journal of Economic Dynamics and Control*, vol. 33, no. 1, pp. 78–92, 2009.
- [68] A. Kononovicius and J. Ruseckas, “Continuous transition from the extensive to the non-extensive statistics in an agent-based herding model,” *European Physics Journal B*, vol. 87, no. 8, p. 169, 2014.